FIFE Workload Management
(JOBSUB + GLIDEINWMS)

PARAG MHASHILKAR
DENNIS BOX

Topics for Discussion

- Job Submission Tools
- JobSub - New Architecture
- GlideinWMS Overview
- Interfacing with JobSub
  - Submitting jobs
  - Fetching job logs
  - Managing jobs
- Statistics
- Current Status & Future Plans
- More Info

What Job Submission Looks Like …

- Typical steps when user wants to run jobs
  - Login in to <experiment>gpvm0x.fnal.gov
  - Setup the environment (krb tickets, proxies etc)
  - Setup the job submission tools (jobsub_tools)
  - Submit jobs using: jobsub
- Common tasks that can be trivialized for the users
  - Access SAM datasets (open/read/close)
  - Use ifdh to transfer data in some cases
  - Run jobs on local clusters, Fermigrid, OSG
  - Limit the number of jobs running at a given time
  - Create and/or submit tarball
  - Limited credential management on behalf of the user

Job Submission Tools

- jobsub_tools
  - Suite of tools to manage batch/grid submissions
  - Simplify job submission process
    - Define common interfaces for difference FIFE Experiments
      - Users associated with difference experiments need to interface with same set of tools
    - Facilitate data I/O using community used tools
    - Integrate complex grid tools sensibly (GlideinWMS, OSG client, …)
    - Protect shared resources from overload (via use of ifdh)
    - Shields the users from changes to the underlying components/ infrastructure

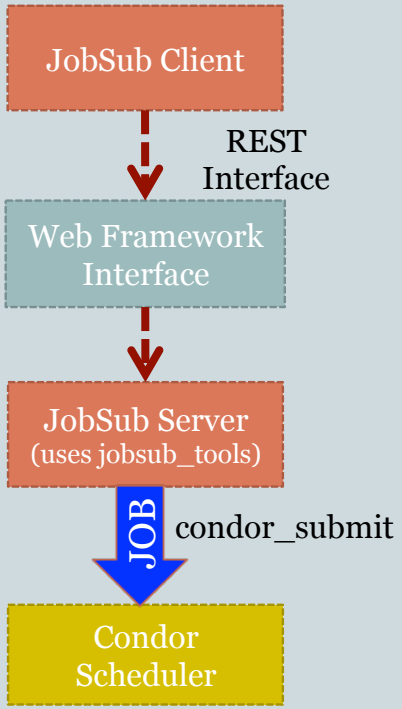
Jobsub_tools: Limitations

- Single point of failure
  - Submission needs schedd and the schedd machine up all the time
  - Deciding between the schedds is user burden
    - One of the schedd is overloaded
    - One of the schedd machine has hardware problems
- No sharing of queuing resources
  - Nodes are experiment specific
- Earlier versions
  - Closely tied to the Fermilab Environment
  - Supporting new experiments always required new jobsub_tools version

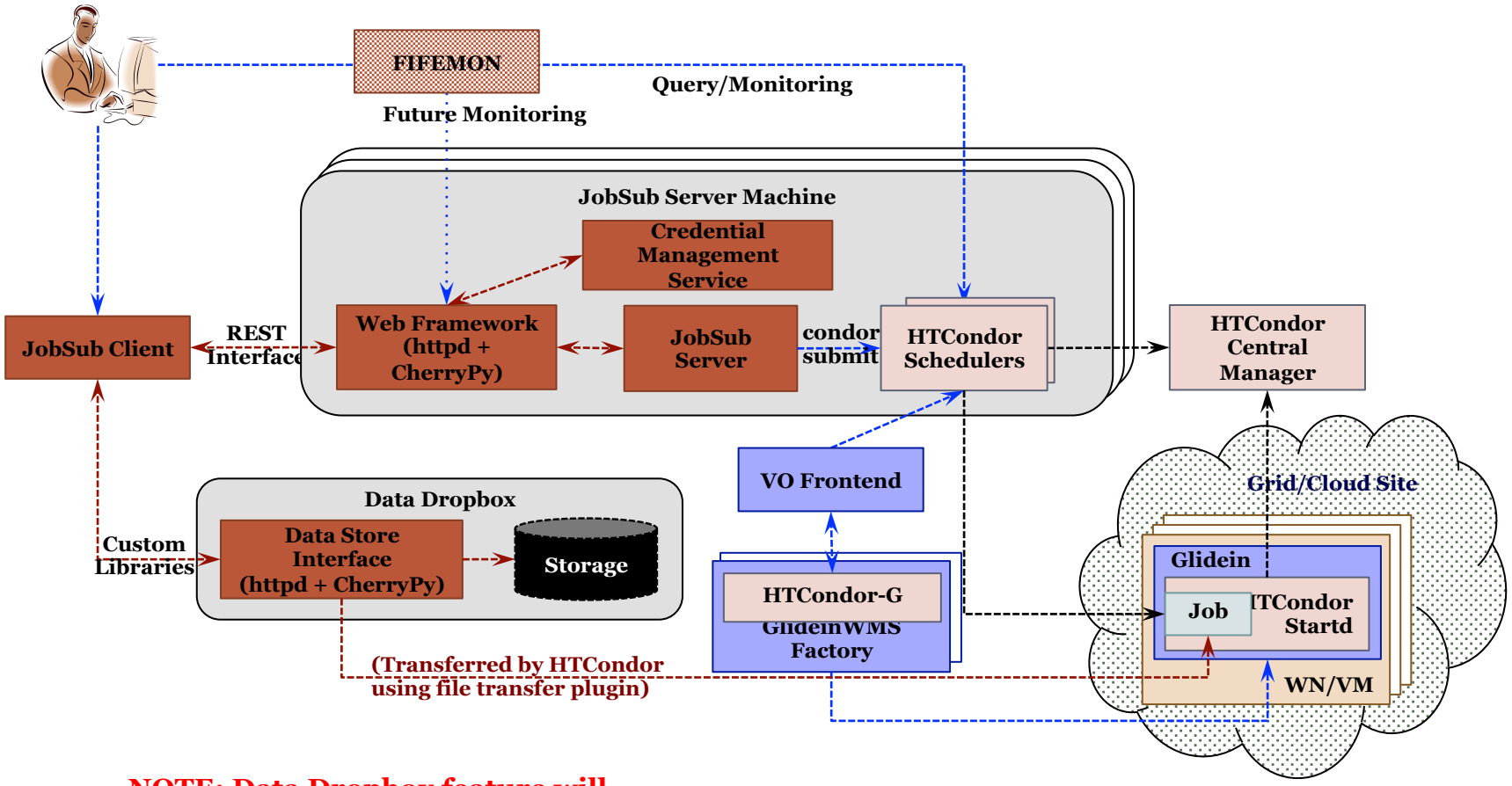
[…]

JobSub - New Architecture

- Architecture Type: Client - Server
  - Thin client & Thick server
  - REST APIs for cleaner design
- Modular design
  - Ease of enhancing/changing client-server communication protocol
  - Web framework can be replaced (Needs rewrite of plumbing code)
- Security
  - Currently X509 with possibility to adapt to the changes in the grid community
- Scalability
  - Multiple servers for load balancing & HA



JobSub: Architecture with GlideinWMS Services



NOTE: Data Dropbox feature will be implemented in future releases.

GlideinWMS

Pilot-based WMS that creates *on demand* a *dynamically-sized overlay condor batch system* on Grid & Cloud resources to address the complex needs of VOs in running application workflows

- GlideinWMS is the resource provision layer
  - Enables submission to different types of resources like Grid & Clouds
  - Responsible for maintaining pressure of idle glideins
- JobSub forms the submission layer and interfaces with the GlideinWMS on behalf of the users
- GlideinWMS components used by FIFE community
  - Configured and maintained by GSCO group
  - Configuration tailored for FIFE

… Would love to talk for several hours on GlideinWMS but …

Check my talk from FIFE 2013 workshop
https://indico.fnal.gov/conferenceDisplay.py?confId=6895

Submitting a Job

- Command: jobsub_submit.py
- Similar to jobsub command in jobsub_tools
- Example: Submit jobs to dedicated slots (with quota)

$ jobsub_submit.py -G nova --resource-provides=usage_model=DEDICATED file://$HOME/nova.sh 1
Server response code: 200
Response OUTPUT:
/fife/local/scratch/uploads/nova/dbox/2014-06-09_144727.804063_7006

/fife/local/scratch/uploads/nova/dbox/2014-06-09_144727.804063_7006/nova.sh_20140609_144728_12037_0_1.cmd

submitting....

Submitting job(s).

1 job(s) submitted to cluster 269.

JobsubJobId of first job: 269.0@fifebatch2.fnal.gov

Use job id 269.0@fifebatch2.fnal.gov to retrieve output

Remote Submission Processing Time: 0.379333019257 sec

Steering Jobs to Resources

- Use --resource-provides to steer jobs to a site or group of sites that satisfy some of the constraints
  - **Example:** --resource-provides=usage_model=<USAGE MODEL>

Resource Type	Usage Model
Amazon	PAID_CLOUD
Fermigrid: Quota'ed Slots	DEDICATED
Fermigrid: Opportunistic Slots	OPPORTUNISTIC
FermiCloud	FERMICLOUD
OSG Sites	OFFSITE

Checking the Job Queue

$ jobsub_q.py

```
Server response code: 200
Response OUTPUT:
170.0@fifebatch2.fnal.gov boyd          06/09       13:09          0+00:00:00 I              0
                0.0             grid_test.sh_20140609_130953_8432_0_1_wrap.sh
224.0@fifebatch2.fnal.gov boyd          06/09       13:56          0+00:00:00 I              0
                0.0             grid_test.sh_20140609_135635_27056_0_1_wrap.sh
226.0@fifebatch2.fnal.gov willis        06/09       14:08          0+00:00:00 I              0
                0.0             grid_test.sh_20140609_140838_30288_0_1_wrap.sh
[...]
261.0@fifebatch2.fnal.gov neha          06/09       14:46          0+00:00:00 I              0                0.0
                testjob-gwms1.sh_20140609_144647_10866_0_1_wrap.sh
262.0@fifebatch2.fnal.gov neha          06/09       14:46          0+00:00:00 I              0                0.0
                testjob-gwms1.sh_20140609_144648_11001_0_1_wrap.sh
Remote Listing Processing Time: 0.108530044556 sec
```

 # QUERY A SINGLE JOB

$ jobsub_q.py -G nova --jobid 269.0@fifebatch2.fnal.gov

```
Server response code: 200
Response OUTPUT:
269.0@fifebatch2.fnal.gov dbox          06/09       14:47          0+00:00:00 I              0
                0.0             nova.sh_20140609_144728_12037_0_1_wrap.sh
Remote Listing Processing Time: 0.0935039520264 sec
```

Fetching Job Output

$ jobsub_fetchlog.py -G nova --jobid 269.0@fifebatch2.fnal.gov

Downloaded to 269.0@fifebatch2.fnal.gov.tgz
Remote Submission Processing Time: 0.239305973053 sec

$ jobsub_fetchlog.py -h
Usage: jobsub_fetchlog.py [options]

Options:
--version show program's version number and exit
-h, --help show this help message and exit
-G <Group/Experiment/Subgroup>, --group=<Group/Experiment/Subgroup>
Group/Experiment/Subgroup for priorities and
accounting
-J <Job ID>, --job=<Job ID>, --jobid=<Job ID>
Job ID
--jobsub-server=<JobSub Server>
Alternate location of JobSub server to use
--timeout=<Timeout> Timeout for the operation in sec
--unzipdir=<Unzip Dir>
Directory to automatically unzip logs into
--archive-format=<Archive Format>
format for downloaded archive:"tar"
(default, compressed) or "zip"

Put Job in Hold State

$jobsub_q.py -G nova
Server response code: 200
Response OUTPUT:
323.0@fifebatch2.fnal.gov dbox 06/09 16:28 0+00:00:00 I 0
0.0 nova.sh_20140609_162840_8980_0_1_wrap.sh
324.0@fifebatch2.fnal.gov dbox 06/09 16:28 0+00:00:00 I 0
0.0 nova.sh_20140609_162843_9114_0_1_wrap.sh
325.0@fifebatch2.fnal.gov dbox 06/09 16:28 0+00:00:00 I 0
0.0 nova.sh_20140609_162847_9246_0_1_wrap.sh
Remote Listing Processing Time: 0.0905869007111 sec
$ jobsub_hold.py -G nova --jobid 323.0@fifebatch2.fnal.gov,324.0@fifebatch2.fnal.gov,325.0@fifebatch2.fnal.gov
Server response code: 200
Response OUTPUT:
Performed Hold on 1 jobs matching your request
Remote Hold Processing Time: 0.0777430534363 sec
[…]
Remote Hold Processing Time: 0.0582039356232 sec
$ jobsub_q.py -G nova
Server response code: 200
Response OUTPUT:
323.0@fifebatch2.fnal.gov dbox 06/09 16:28 0+00:00:00 H 0
0.0 nova.sh_20140609_162840_8980_0_1_wrap.sh
324.0@fifebatch2.fnal.gov dbox 06/09 16:28 0+00:00:00 H 0
0.0 nova.sh_20140609_162843_9114_0_1_wrap.sh
325.0@fifebatch2.fnal.gov dbox 06/09 16:28 0+00:00:00 H 0
0.0 nova.sh_20140609_162847_9246_0_1_wrap.sh
Remote Listing Processing Time: 0.0905029773712 sec

Releasing a Job

$ jobsub_q.py -G nova
Server response code: 200
Response OUTPUT:
323.0@fifebatch2.fnal.gov dbox 06/09 16:28 0+00:00:00 H 0
0.0 nova.sh_20140609_162840_8980_0_1_wrap.sh
324.0@fifebatch2.fnal.gov dbox 06/09 16:28 0+00:00:00 H 0
0.0 nova.sh_20140609_162843_9114_0_1_wrap.sh
325.0@fifebatch2.fnal.gov dbox 06/09 16:28 0+00:00:00 H 0
0.0 nova.sh_20140609_162847_9246_0_1_wrap.sh
Remote Listing Processing Time: 0.0905029773712 sec
$ jobsub_release.py -G nova --jobid 323.0@fifebatch2.fnal.gov,324.0@fifebatch2.fnal.gov, 325.0@fifebatch2.fnal.gov
Server response code: 200
Response OUTPUT:
Performed Release on 1 jobs matching your request
Remote Release Processing Time: 0.0870110988617 sec
[...]
Remote Release Processing Time: 0.0682098865509 sec
$ jobsub_q.py -G nova
Server response code: 200
Response OUTPUT:
323.0@fifebatch2.fnal.gov dbox 06/09 16:28 0+00:00:00 I 0
0.0 nova.sh_20140609_162840_8980_0_1_wrap.sh
324.0@fifebatch2.fnal.gov dbox 06/09 16:28 0+00:00:00 I 0
0.0 nova.sh_20140609_162843_9114_0_1_wrap.sh
325.0@fifebatch2.fnal.gov dbox 06/09 16:28 0+00:00:00 I 0
0.0 nova.sh_20140609_162847_9246_0_1_wrap.sh
Remote Listing Processing Time: 0.0906929969788 sec

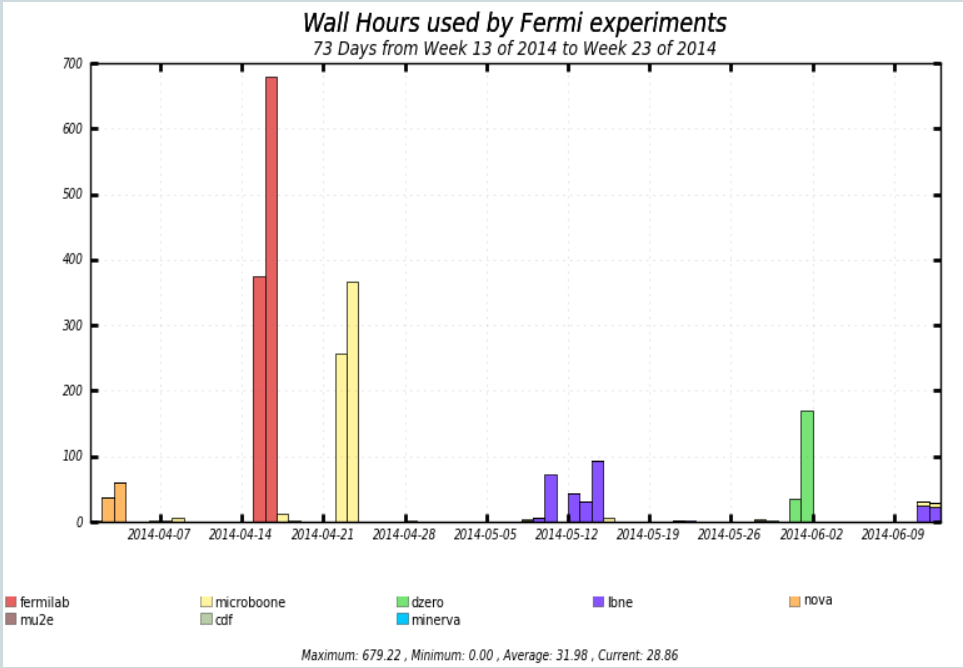
Removing Jobs

$ jobsub_q.py -G nova
Server response code: 200
Response OUTPUT:
323.0@fifebatch2.fnal.gov dbox 06/09 16:28 0+00:00:00 H 0
0.0 nova.sh_20140609_162840_8980_0_1_wrap.sh
324.0@fifebatch2.fnal.gov dbox 06/09 16:28 0+00:00:00 H 0
0.0 nova.sh_20140609_162843_9114_0_1_wrap.sh
325.0@fifebatch2.fnal.gov dbox 06/09 16:28 0+00:00:00 H 0
0.0 nova.sh_20140609_162847_9246_0_1_wrap.sh
Remote Listing Processing Time: 0.0907280445099 sec

$ jobsub_rm.py -G nova --jobid 323.0@fifebatch2.fnal.gov,324.0@fifebatch2.fnal.gov, 325.0@fifebatch2.fnal.gov
Server response code: 200
Response OUTPUT:
Performed Remove on 1 jobs matching your request
Remote Removal Processing Time: 0.0708208084106 sec
Server response code: 200
Response OUTPUT:
Performed Remove on 1 jobs matching your request
Remote Removal Processing Time: 0.0634789466858 sec
Server response code: 200
Response OUTPUT:
Performed Remove on 1 jobs matching your request
Remote Removal Processing Time: 0.0582299232483 sec

Usage Plots

- Process for on boarding the experiments
- JobSub Group & GCSO work closely with the experiment users
  - Configure the GlideinWMS to support new experiment
  - User submits sample jobs
  - Work with the sites to debug issues
  - Gradually move to production



Current Status & Future Plans

- Current Version: v0.3.1.2
  - Supports HA (v0.3 series)
  - Deployed on the fifebatch.fnal.gov & the pre-prod servers
  - Release Notes: https://cdcvs.fnal.gov/redmine/projects/jobsub/wiki/Release_Notes
- Next Major Series: v0.4 (~3rd week of July)
  - Identify and address critical missing features
  - Support for Run II experiments (CDF/Dzero)
- First Stable Release: v1.0
  - Last release from v0.4 series to be released as first stable version if no major show stoppers are identified
- More info on releases and the process https://cdcvs.fnal.gov/redmine/projects/jobsub/wiki#JobSub-Releases

Documentation

- JobSub
  - Project Homepage
    https://cdcvs.fnal.gov/redmine/projects/jobsub/wiki
  - Users guide
    https://cdcvs.fnal.gov/redmine/projects/jobsub/wiki/Using_the_Client
  - Issue Tracker
    https://cdcvs.fnal.gov/redmine/projects/jobsub/issues
- GlideinWMS
  - Project Homepage
    http://www.uscms.org/SoftwareComputing/Grid/WMS/glideinWMS
  - Issue Tracker
    https://cdcvs.fnal.gov/redmine/projects/glideinwms/issues